How to use interaction terms in BMA: Reply to Crespo's comment on Masanjala and Papageorgiou (2008)

Chris Papageorgiou
IMF
E-mail: cpapageorgiou@imf.org

June 3, 2011

Abstract. Jesus Crespo Cuaresma (forthcoming) shows that the results in Masanjala and Papageorgiou (Rough and lonely road to prosperity: a reexamination of the sources of growth in Africa using Bayesian model averaging, *Journal of Applied Econometrics* 2008; 23(5): 671-682) are sensitive to an alternative prior model structure in considering interaction terms. As a side-issue, the algorithm for averaging over models is also challenged. In this reply I show that the prior used in Masanjala and Papageorgiou is as sensible as the prior suggested by Crespo. What we learn from Crespo's comment and this reply is that further effort should be dedicated to the study of parameter heterogeneity in the framework of BMA methods.

Jesus Crespo Cuaresma has written a thoughtful comment on Masanjala and Papageorgiou (2008; henceforth MP) pointing out that the strong heterogeneity of growth determinants for Sub-saharan Africa found in that paper could be a spurious result based on the wrong specification of the prior over the model space. Specifically, it is argued that the way MP incorporated interaction terms in their model averaging exercise was not appropriate; when using an alternative treatment of interaction terms the heterogeneity results are once again eliminated. In addition, as a side-issue the computational algorithm for conducting the inference was also questioned.

Crespo's main point on the appropriate treatment of interaction terms in BMA is well taken. He argues that using the model space that includes interactions without forcing the interacting regressors to also be present (as was done in MP) is not appropriate as, it biases the results in favor of positive interaction effects. His point is that a model in which there is an interaction should also include all relevant uninteracted regressors. Crespo subsequently shows that when using an alternative model prior which basically does not allow for "orphan" uninteracted terms in any model considered in BMA then the heterogeneity result about Africa is eliminated.

In what follows I argue that MP's choice of using the entire model space (including uninteracted and interacted variables) is quite reasonable. I briefly elaborate below using a few simple equations. It should be noted that the following argumentation is entirely due to Eduardo Ley with whom I had numerous stimulating discussions on the subject over several months.[1]

[1] I also thank Adrian Raftery for enriching these discussions.

Assume that d_i is a categorical variable that takes the value 1 when $i \in I$ (e.g., country belongs to Africa) and 0 otherwise. The full model takes the form:

$$y_i = \alpha + \beta x_i + \gamma d_i x_i + \varepsilon_i, \tag{1}$$

which implies:

$$\begin{aligned} E[y_i|x_i, d_i &= 1] = \alpha + (\beta + \gamma)x_i \\ E[y_i|x_i, d_i &= 0] = \alpha + \beta x_i. \end{aligned}$$

The model without x_i is given by:

$$y_i = \alpha + \gamma d_i x_i + \varepsilon_i. \tag{2}$$

Now we have

$$\begin{aligned} E[y_i|x_i, d_i &= 1] = \alpha + \gamma x_i \\ E[y_i|x_i, d_i &= 0] = \alpha. \end{aligned}$$

The main question for this setup is whether we can 'average' (1) and (2)? I argue that there is no reason to weigh against this. In both cases, γ picks up the extra effect that x has on y when $i \in I$. The fact that in (1) this *extra* effect is with respect to β and in (2) it is with respect to $\beta = 0$ is not troublesome. In either case it represents the I dividend. Equivalently, the model without $x_i d_i$ is given by:

$$y_i = \alpha + \beta x_i + \varepsilon_i. \tag{3}$$

We now have:

$$E[y_i|x_i, d_i = 1, 0] = \alpha + \beta x_i.$$

I see no obvious reservation in interpreting β across (1)–(3), since in all cases it picks up the overall response of y to x, when no special attention is paid to I. This simple setup shows that MP's prior is quite reasonable. Furthermore, this simple argumentation makes the case that the question about the treatment of interaction terms in BMA is really a question about model prior setting, a view shared by Crespo who considers Chipman's (1996) strong heredity prior over the model space that could "more appropriate" deal with parameter heterogeneity across subsamples.

In that case what is needed is a comparison of prior structures using out-of-sample prediction, precisely what Crespo attempted to do using log predictive score (LPS) to measure predictive accuracy. He has shown that replicating the procedure for 100 random out-of-sample partitions obtained an average LPS of 3.21 for the MP prior and 3.30 for the strong heredity prior, arguing evidence against using the MP prior. Although, I am in compete agreement about using predictive performance as it is a natural and neutral criterion for comparing different priors, I do not find it particularly convincing that a 0.09 points difference in LPS score (especially when the procedure was replicated for only 100 partitions) is sufficient to argue against the MP prior. It would be useful to know what is the LPS median and standard deviation; the later should be informative for judging the significance of the 0.09 value. Eicher, Papageorgiou and Raftery (forthcoming) demonstrated that LPS can be quite sensitive to both the number of partitions used (sometimes requiring several hundreds of partitions before LPS can discriminate between priors) and data outliers (likely in our sample); these authors have used an alternative criterion, namely the continuous ranked probability score (CRPS; Matheson and Winkler, 1976) shown to be less sensitive to outliers. Therefore, a more comprehensive assessment of predictive accuracy of the two alternative priors is required before arriving at firm conclusions.

Regarding the secondary point about the choice of computational algorithms Crespo states: "It should be noted that the BMA literature often warns that for such a large set of covariates as the one presented in MP, the *leaps and bounds* method is known to explore regions of the model space which may not be sufficiently large enough to obtain reliable estimates (see, for example, Clyde, 1999, 2001)." Although I fully recognize the warnings of the literature, Crespo's comment seems to prematurely conclude that MC^3 should be universally chosen over the *leaps and bounds* algorithm. In my view future work which employs a rather straight forward Monte Carlo exercise, using simulated and factual data, could usefully clarify this issue.

In summary, I thank Crespo for his thoughtful comment which enhances our understanding of BMA methods. I fully agree that MP's results are sensitive to the choice of prior and the computational method, but as Crespo states ".. it is not surprising that reasonable changes in the prior over the model space result in different inferences concerning the importance of parameter heterogeneity in the sample at hand." I further argue that the prior structure assumed in MP is in principle as reasonable as the strong heredity prior favored by Crespo. A careful and comprehensive out-of-sample prediction exercise based on cross-validated predictive performance, as in Ley and

Steel (2009) and Eicher, Papageorgiou and Raftery (forthcoming), is warranted to properly evaluate which prior is more appropriate for the application at hand. While the original paper has created a fruitful discussion which has moved the research frontier in BMA applications forward, its main finding, that there is large heterogeneity in Africa, is still open to future developments in BMA analysis and better quality and more complete data for the region.

References

[1] Chipman H.A. (1996)."Bayesian variable selection with related predictors,". *Canadian Journal of Statistics* 24: 17-36.

[2] Clyde M. (1999). "Bayesian model averaging and model search strategies," *Bayesian Statistics* 6: 157-185.

[3] Clyde M. (2001). "Discussion of 'The impractical implementation of Bayesian model selection"'. in IMS Lecture Notes: Monograph Series, Vol. 38. Institute of Mathematical Statistics: Beachwood, OH; 117-124.

[4] Crespo, J.C. (forthcoming). "How different is Africa? A comment on Masanjala and Papageorgiou" *Journal of Applied Econometrics*.

[5] Eicher, T.S., Papageorgiou, C., Raftery, A.E. (forthcoming). "Default Priors and Predictive Performance in Bayesian Model Averaging, with Application to Growth Determinants," *Journal of Applied Econometrics*.

[6] Ley, E. and Steel, M.F.J. (2009). "On the Effect of Prior Assumptions in BMA with Applications to Growth Regression," *Journal of Applied Econometrics,* 24: 651-674

[7] Masanjala, W. and Papageorgiou, C. (2008). "Rough and lonely road to prosperity: a reexamination of the sources of growth in Africa using Bayesian model averaging," *Journal of Applied Econometrics*, 23: 671-682.

[8] Sala-i-Martin, X., Doppelhofer, G. and Miller, R.I. (2004). "Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach," *American Economic Review*, 94: 813-835.

[9] Raftery, A.E., Madigan, D. and Volinsky, C.T. (1995). "Accounting for model uncertainty in survival analysis improves predictive performance (with Discussion)," in *Bayesian Statistics 5* (J.M. Bernardo, J.O. Berger, A.P. Dawid and A.F.M. Smith, eds.), Oxford University Press, 323-349.